

#2

OIPE

RAW SEQUENCE LISTING

DATE: 12/14/2001

PATENT APPLICATION: US/10/005,647

TIME: 10:25:08

Input Set : A:\seqlist_1708 0240001.txt

Output Set: N:\CRF3\12142001\I005647.raw

ENTERED

5 <110> APPLICANT: Jalkanen, Markku
 7 Darwish, Kamel El
 9 Lindahl, Ulf
 11 Li, Jin-Ping
 15 <120> TITLE OF INVENTION: Glucuronyl C5-Epimerase, DNA Encoding the Same and Uses
 Thereof
 19 <130> FILE REFERENCE: 1708.0280002
 C--> 21 <140> CURRENT APPLICATION NUMBER: US/10/005,647
 C--> 23 <141> CURRENT FILING DATE: 2001-12-07
 26 <150> PRIOR APPLICATION NUMBER: US 60/304,180
 28 <151> PRIOR FILING DATE: 2000-12-08
 31 <150> PRIOR APPLICATION NUMBER: US 09/732,026
 33 <151> PRIOR FILING DATE: 2000-12-08
 37 <160> NUMBER OF SEQ ID NOS: 2
 41 <170> SOFTWARE: PatentIn version 3.0
 45 <210> SEQ ID NO: 1
 47 <211> LENGTH: 1854
 49 <212> TYPE: DNA
 51 <213> ORGANISM: Mus musculus
 55 <220> FEATURE:
 57 <221> NAME/KEY: CDS
 59 <222> LOCATION: (1)..(1854)
 63 <400> SEQUENCE: 1
 64 atg cgt tgt ttg gca gct cgg gtc aac tat aag act ttg att atc atc 48
 65 Met Arg Cys Leu Ala Ala Arg Val Asn Tyr Lys Thr Leu Ile Ile Ile
 66 1 5 10 15
 68 tgt gcg cta ttc act ttg gtc aca gta ctt ttg tgg aat aag tgt tcc 96
 69 Cys Ala Leu Phe Thr Leu Val Thr Val Leu Leu Trp Asn Lys Cys Ser
 70 20 25 30
 72 agc gac aaa gca atc cag ttt cct cgg cac ttg agt agt gga ttc aga 144
 73 Ser Asp Lys Ala Ile Gln Phe Pro Arg His Leu Ser Ser Gly Phe Arg
 74 35 40 45
 76 gtg gat gga tta gaa aaa aga tca gca gca tct gaa agt aac cac tat 192
 77 Val Asp Gly Leu Glu Lys Arg Ser Ala Ala Ser Glu Ser Asn His Tyr
 78 50 55 60
 80 gcc aac cac ata gcc aaa cag cag tca gaa gag gca ttt cct cag gaa 240
 81 Ala Asn His Ile Ala Lys Gln Gln Ser Glu Glu Ala Phe Pro Gln Glu
 82 65 70 75 80
 84 caa cag aag gca ccc cct gtt gtt ggg ggc ttc aat agc aac ggg gga 288
 85 Gln Gln Lys Ala Pro Pro Val Val Gly Gly Phe Asn Ser Asn Gly Gly
 86 85 90 95
 88 agc aag gtg tta ggg ctc aaa tat gaa gag att gac tgt ctc ata aac 336
 89 Ser Lys Val Leu Gly Leu Lys Tyr Glu Glu Ile Asp Cys Leu Ile Asn
 90 100 105 110
 92 gat gag cac acc att aaa ggg aga cga gag ggg aat gaa gtt ttc ctt 384
 93 Asp Glu His Thr Ile Lys Gly Arg Glu Gly Asn Glu Val Phe Leu
 94 115 120 125
 96 cca ttc act tgg gta gag aaa tac ttt gat gtt tat gga aaa gtg gtc 432

RAW SEQUENCE LISTING

DATE: 12/14/2001

PATENT APPLICATION: US/10/005,647

TIME: 10:25:08

Input Set : A:\seqlist_1708 0240001.txt

Output Set: N:\CRF3\12142001\I005647.raw

```

97 Pro Phe Thr Trp Val Glu Lys Tyr Phe Asp Val Tyr Gly Lys Val Val
98      130      135      140
100 cag tat gac ggc tat gat cga ttt gaa ttc tct cat agc tat tcc aaa      480
101 Gln Tyr Asp Gly Tyr Asp Arg Phe Glu Phe Ser His Ser Tyr Ser Lys
102 145      150      155      160
104 gtc tat gca cag aga tca cct tat cac cct gac ggt gtg ttt atg tcc      528
105 Val Tyr Ala Gln Arg Ser Pro Tyr His Pro Asp Gly Val Phe Met Ser
106      165      170      175
108 ttt gaa ggc tac aat gtg gaa gtc cga gac aga gtc aaa tgt ata agt      576
109 Phe Glu Gly Tyr Asn Val Glu Val Arg Asp Arg Val Lys Cys Ile Ser
110      180      185      190
112 gga gtt gaa ggt gtg cca tta tct acc cag tgg ggg cct caa ggc tat      624
113 Gly Val Glu Gly Val Pro Leu Ser Thr Gln Trp Gly Pro Gln Gly Tyr
114      195      200      205
116 ttc tac cca atc cag att gca cag tat ggg cta agt cat tac agc aag      672
117 Phe Tyr Pro Ile Gln Ile Ala Gln Tyr Gly Leu Ser His Tyr Ser Lys
118      210      215      220
120 aat cta acc gag aaa ccc cct cac ata gaa gta tat gaa aca gca gaa      720
121 Asn Leu Thr Glu Lys Pro Pro His Ile Glu Val Tyr Glu Thr Ala Glu
122 225      230      235      240
124 gac agg gac aga aac atc aga cct aat gaa tgg act gtg ccc aag ggg      768
125 Asp Arg Asp Arg Asn Ile Arg Pro Asn Glu Trp Thr Val Pro Lys Gly
126      245      250      255
128 tgc ttc atg gcc agt gtg gca gac aag tct aga tcc acc aat gtt aaa      816
129 Cys Phe Met Ala Ser Val Ala Asp Lys Ser Arg Ser Thr Asn Val Lys
130      260      265      270
132 cag ttt att gct cca gaa acc agt gaa ggt gtg tct ttg cag ctg gga      864
133 Gln Phe Ile Ala Pro Glu Thr Ser Glu Gly Val Ser Leu Gln Leu Gly
134      275      280      285
136 aac aca aaa gac ttc att att tca ttt gac ctc aag ctt tta aca aat      912
137 Asn Thr Lys Asp Phe Ile Ile Ser Phe Asp Leu Lys Leu Leu Thr Asn
138      290      295      300
140 ggg agt gtg tct gtg gtt ctg gag acc aca gaa aag aat cag ctc ttc      960
141 Gly Ser Val Ser Val Val Leu Glu Thr Thr Glu Lys Asn Gln Leu Phe
142 305      310      315      320
144 act gtg cat tat gtc tca aac acc cag ctg att gct ttc aga gac agg      1008
145 Thr Val His Tyr Val Ser Asn Thr Gln Leu Ile Ala Phe Arg Asp Arg
146      325      330      335
148 gac ata tac tac ggc att ggg ccc aga act tca tgg agt aca gtt acc      1056
149 Asp Ile Tyr Tyr Gly Ile Gly Pro Arg Thr Ser Trp Ser Thr Val Thr
150      340      345      350
152 aga gac ctg gtc act gac ctc agg aaa gga gtg ggc ctt tct aac aca      1104
153 Arg Asp Leu Val Thr Asp Leu Arg Lys Gly Val Gly Leu Ser Asn Thr
154      355      360      365
156 aaa gct gtc aag cca acc aaa atc atg ccc aaa aag gtg gtt agg ttg      1152
157 Lys Ala Val Lys Pro Thr Lys Ile Met Pro Lys Lys Val Val Arg Leu
158      370      375      380
160 att gca aaa ggg aag gga ttc ctg gac aac att acc atc tca acc aca      1200
161 Ile Ala Lys Gly Lys Gly Phe Leu Asp Asn Ile Thr Ile Ser Thr Thr

```

RAW SEQUENCE LISTING

DATE: 12/14/2001

PATENT APPLICATION: US/10/005,647

TIME: 10:25:08

Input Set : A:\seqlist_1708 0240001.txt

Output Set: N:\CRF3\12142001\I005647.raw

```

162 385          390          395          400
164 gcc cac atg gct gca ttc ttt gct gca agt gac tgg cta gtg agg aac 1248
165 Ala His Met Ala Ala Phe Phe Ala Ala Ser Asp Trp Leu Val Arg Asn
166          405          410          415
168 cag gat gag aaa ggt ggc tgg cca att atg gtg acc cgg aag tta ggg 1296
169 Gln Asp Glu Lys Gly Gly Trp Pro Ile Met Val Thr Arg Lys Leu Gly
170          420          425          430
172 gaa ggg ttt aaa tct tta gaa cca gga tgg tac tct gcc atg gca caa 1344
173 Glu Gly Phe Lys Ser Leu Glu Pro Gly Trp Tyr Ser Ala Met Ala Gln
174          435          440          445
176 ggg caa gcc atc tct acc tta gtc agg gcc tat ctt cta acg aaa gac 1392
177 Gly Gln Ala Ile Ser Thr Leu Val Arg Ala Tyr Leu Leu Thr Lys Asp
178          450          455          460
180 tat gta ttc ctc agt tca gct tta agg gca aca gcc cca tac aag ttt 1440
181 Tyr Val Phe Leu Ser Ser Ala Leu Arg Ala Thr Ala Pro Tyr Lys Phe
182 465          470          475          480
184 ccg tca gag cag cat gga gtt aaa gcc gtg ttc atg aat aaa cat gac 1488
185 Pro Ser Glu Gln His Gly Val Lys Ala Val Phe Met Asn Lys His Asp
186          485          490          495
188 tgg tat gaa gaa tat cca acc aca cct agc tct ttt gtt tta aat ggc 1536
189 Trp Tyr Glu Glu Tyr Pro Thr Thr Pro Ser Ser Phe Val Leu Asn Gly
190          500          505          510
192 ttt atg tat tct tta att ggg ctg tat gac cta aaa gaa aca gca ggg 1584
193 Phe Met Tyr Ser Leu Ile Gly Leu Tyr Asp Leu Lys Glu Thr Ala Gly
194          515          520          525
196 gag aca ctt ggg aaa gaa gca agg tcc ttg tac gag cgc ggc atg gaa 1632
197 Glu Thr Leu Gly Lys Glu Ala Arg Ser Leu Tyr Glu Arg Gly Met Glu
198          530          535          540
200 tct ctt aaa gcc atg ctg ccc ttg tat gat act ggc tcc ggg acc atc 1680
201 Ser Leu Lys Ala Met Leu Pro Leu Tyr Asp Thr Gly Ser Gly Thr Ile
202 545          550          555          560
204 tat gac ctc cgc cac ttc atg ctt ggc att gct ccc aac ctg gcc cgc 1728
205 Tyr Asp Leu Arg His Phe Met Leu Gly Ile Ala Pro Asn Leu Ala Arg
206          565          570          575
208 tgg gac tat cac acc acc cac att aac cag ctg cag ctg ctc agc acc 1776
209 Trp Asp Tyr His Thr Thr His Ile Asn Gln Leu Gln Leu Leu Ser Thr
210          580          585          590
212 atc gat gag tcc cca atc ttc aaa gaa ttt gtc aag agg tgg aaa agc 1824
213 Ile Asp Glu Ser Pro Ile Phe Lys Glu Phe Val Lys Arg Trp Lys Ser
214          595          600          605
216 tac ctt aaa ggc agt agg gca aag cac aac 1854
217 Tyr Leu Lys Gly Ser Arg Ala Lys His Asn
218          610          615
221 <210> SEQ ID NO: 2
223 <211> LENGTH: 618
225 <212> TYPE: PRT
227 <213> ORGANISM: Mus musculus
231 <400> SEQUENCE: 2
233 Met Arg Cys Leu Ala Ala Arg Val Asn Tyr Lys Thr Leu Ile Ile Ile

```

RAW SEQUENCE LISTING

DATE: 12/14/2001

PATENT APPLICATION: US/10/005,647

TIME: 10:25:08

Input Set : A:\seqlist_1708 0240001.txt

Output Set: N:\CRF3\12142001\I005647.raw

```

234 1          5          10          15
237 Cys Ala Leu Phe Thr Leu Val Thr Val Leu Leu Trp Asn Lys Cys Ser
238          20          25          30
241 Ser Asp Lys Ala Ile Gln Phe Pro Arg His Leu Ser Ser Gly Phe Arg
242          35          40          45
245 Val Asp Gly Leu Glu Lys Arg Ser Ala Ala Ser Glu Ser Asn His Tyr
246          50          55          60
249 Ala Asn His Ile Ala Lys Gln Gln Ser Glu Glu Ala Phe Pro Gln Glu
250 65          70          75          80
253 Gln Gln Lys Ala Pro Pro Val Val Gly Gly Phe Asn Ser Asn Gly Gly
254          85          90          95
257 Ser Lys Val Leu Gly Leu Lys Tyr Glu Glu Ile Asp Cys Leu Ile Asn
258          100          105          110
261 Asp Glu His Thr Ile Lys Gly Arg Arg Glu Gly Asn Glu Val Phe Leu
262          115          120          125
265 Pro Phe Thr Trp Val Glu Lys Tyr Phe Asp Val Tyr Gly Lys Val Val
266          130          135          140
269 Gln Tyr Asp Gly Tyr Asp Arg Phe Glu Phe Ser His Ser Tyr Ser Lys
270 145          150          155          160
273 Val Tyr Ala Gln Arg Ser Pro Tyr His Pro Asp Gly Val Phe Met Ser
274          165          170          175
277 Phe Glu Gly Tyr Asn Val Glu Val Arg Asp Arg Val Lys Cys Ile Ser
278          180          185          190
281 Gly Val Glu Gly Val Pro Leu Ser Thr Gln Trp Gly Pro Gln Gly Tyr
282          195          200          205
285 Phe Tyr Pro Ile Gln Ile Ala Gln Tyr Gly Leu Ser His Tyr Ser Lys
286          210          215          220
289 Asn Leu Thr Glu Lys Pro Pro His Ile Glu Val Tyr Glu Thr Ala Glu
290 225          230          235          240
293 Asp Arg Asp Arg Asn Ile Arg Pro Asn Glu Trp Thr Val Pro Lys Gly
294          245          250          255
297 Cys Phe Met Ala Ser Val Ala Asp Lys Ser Arg Ser Thr Asn Val Lys
298          260          265          270
301 Gln Phe Ile Ala Pro Glu Thr Ser Glu Gly Val Ser Leu Gln Leu Gly
302          275          280          285
305 Asn Thr Lys Asp Phe Ile Ile Ser Phe Asp Leu Lys Leu Leu Thr Asn
306          290          295          300
309 Gly Ser Val Ser Val Val Leu Glu Thr Thr Glu Lys Asn Gln Leu Phe
310 305          310          315          320
313 Thr Val His Tyr Val Ser Asn Thr Gln Leu Ile Ala Phe Arg Asp Arg
314          325          330          335
317 Asp Ile Tyr Tyr Gly Ile Gly Pro Arg Thr Ser Trp Ser Thr Val Thr
318          340          345          350
321 Arg Asp Leu Val Thr Asp Leu Arg Lys Gly Val Gly Leu Ser Asn Thr
322          355          360          365
325 Lys Ala Val Lys Pro Thr Lys Ile Met Pro Lys Lys Val Val Arg Leu
326          370          375          380
329 Ile Ala Lys Gly Lys Gly Phe Leu Asp Asn Ile Thr Ile Ser Thr Thr
330 385          390          395          400

```

RAW SEQUENCE LISTING

DATE: 12/14/2001

PATENT APPLICATION: US/10/005,647

TIME: 10:25:08

Input Set : A:\seqlist_1708 0240001.txt

Output Set: N:\CRF3\12142001\I005647.raw

```

333 Ala His Met Ala Ala Phe Phe Ala Ala Ser Asp Trp Leu Val Arg Asn
334          405          410          415
337 Gln Asp Glu Lys Gly Gly Trp Pro Ile Met Val Thr Arg Lys Leu Gly
338          420          425          430
341 Glu Gly Phe Lys Ser Leu Glu Pro Gly Trp Tyr Ser Ala Met Ala Gln
342          435          440          445
345 Gly Gln Ala Ile Ser Thr Leu Val Arg Ala Tyr Leu Leu Thr Lys Asp
346          450          455          460
349 Tyr Val Phe Leu Ser Ser Ala Leu Arg Ala Thr Ala Pro Tyr Lys Phe
350 465          470          475          480
353 Pro Ser Glu Gln His Gly Val Lys Ala Val Phe Met Asn Lys His Asp
354          485          490          495
357 Trp Tyr Glu Glu Tyr Pro Thr Thr Pro Ser Ser Phe Val Leu Asn Gly
358          500          505          510
361 Phe Met Tyr Ser Leu Ile Gly Leu Tyr Asp Leu Lys Glu Thr Ala Gly
362          515          520          525
365 Glu Thr Leu Gly Lys Glu Ala Arg Ser Leu Tyr Glu Arg Gly Met Glu
366          530          535          540
369 Ser Leu Lys Ala Met Leu Pro Leu Tyr Asp Thr Gly Ser Gly Thr Ile
370 545          550          555          560
373 Tyr Asp Leu Arg His Phe Met Leu Gly Ile Ala Pro Asn Leu Ala Arg
374          565          570          575
377 Trp Asp Tyr His Thr Thr His Ile Asn Gln Leu Gln Leu Leu Ser Thr
378          580          585          590
381 Ile Asp Glu Ser Pro Ile Phe Lys Glu Phe Val Lys Arg Trp Lys Ser
382          595          600          605
385 Tyr Leu Lys Gly Ser Arg Ala Lys His Asn
386          610          615

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/005,647

DATE: 12/14/2001

TIME: 10:25:09

Input Set : A:\seqlist_1708 0240001.txt

Output Set: N:\CRF3\12142001\I005647.raw

L:21 M:270 C: Current Application Number differs, Replaced Application Number
L:23 M:271 C: Current Filing Date differs, Replaced Current Filing Date